AN ICA BASED ALGORITHM FOR VIDEO WATERMARKING

Hussein Joumaa *Franck Davoine*

Heudiasyc Laboratory, CNRS, University of Technology of Compiègne
BP 20529, 60205 COMPIEGNE Cedex, FRANCE
{Hussein.Joumaa, Franck.Davoine}@hds.utc.fr

ABSTRACT

In this paper, a new video watermarking scheme is proposed. We adapt two watermarking algorithms, originally proposed for still image watermarking, to embed data in a set of statistically independent sources, extracted from a video sequence. We show the interest of applying such an approach on video watermarking. The proposed scheme offers a good robustness against MPEG compression attack, as well as an important capacity level. We consider in this paper data hiding in digital TV channels where data are compressed using MPEG-2. The main contribution of our study is to compare two different watermarking approaches applied on error prediction frames transformed by ICA.

1. INTRODUCTION

The rapid expansion of the internet and the overall development of digital technologies in the past years have sharply increased the availability of multimedia content. Watermarking, which allows for the secret embedding of information in a host data, has emerged as a widely approach for copyright protection, ownership identification. However, other applications can be considered. In this paper, we are particulary interested in applications requiring a large amount of data hiding. For instance, they include embedded control to track the use of a particular video clip in pay-per-view applications, hidden communications, smart images/video that can self-correct under intentional attacks, to mention a few. The capability to hide large amounts of data will also enable robust hiding of digital watermarks by introducing redundancies in the data.

Techniques for hiding watermarks in still images have grown steadily more sophisticated and increasingly robust to lossy image compression and standard image processing operations, as well as to cryptographic attack. Many of the current techniques for embedding marks in digital images, inspired by methods of image coding and compression, work in a transformed domain. In the case of video watermarking, the challenge is to mark a group of images which are strongly intercorrelated and often manipulated in a compressed form, e.g. MPEG. A first group of watermarking methods therefore directly operates on MPEG data to avoid full decompression [1]. Other researchers watermark MPEG-2 motion compensation vectors [2], or MPEG-4 video objects [3]. In order for the watermark to be less dependent on the way the video compression was done, another approach is to mark the uncompressed video sequence in spite of the increased computational cost. To work with uncompressed video, a possibility is to individually mark all the frames of the video using still image watermarking techniques. We have adopted this approach in our scheme, furthermore we have taken into account the temporal dimension of the video: the watermark has been embedded in error prediction frames. Therefore, we have adapted a method initially used to watermark still images. It is based on embedding the message in a set of statistically independent sources, obtained by independent component analysis (ICA). These sources constitute the spanning of a feature space, and represent the convertext in conjunction with the corresponding set of constant mixing matrices.

The paper is organized as follows : In section 2, we give essential notions about ICA techniques, then we review some proposed watermarking methods based on ICA. In section 3, our watermarking algorithm is detailed. In section 4, experimental results are presented. Finally, conclusions are drawn in section 5.

2. ICA AND WATERMARKING

In this section, we begin by introducing the ICA method. We also review some watermarking methods, based on ICA. ICA is a statistical technique that aims at finding linear projections of a signal that maximize their mutual independence. Its main applications are in feature extraction, and blind source separation [4].

In the watermarking problem, ICA was used by González-Serrano et al. [5]. Their approach has been applied to still image watermarking. It consists in projecting the image into a basis defined by an ICA algorithm. The watermark is then introduced by changing the least significant components of the representation in the new basis. This method was applied for the fragile watermarking problem.

Another approach was proposed by Bounkong et al. [6]. It has been shown that the ICA allows the maximization of the information content and minimization of the induced distortion by decomposing the host signal into statistically independent sources. Embedding information in one of these sources minimizes the emerging cross-channel interference. In fact for a broadcast class of attacks, and fixed capacity values, it has been shown that distortion is minimized when the message is embedded in statistically independent sources [6]. Information analysis also proves that the information hiding capacity of statistically independent sources is maximal [7].
We are interested in the change of basis provided by ICA. Consider a $k^2 \times 1$ vector $\mathbf{x}_t$ projected into a space of p components $\mathbf{y}_t$ as statistically independent as possible. The ICA problem consists of finding this change of basis represented by a $p \times k^2$ matrix A, which is called the mixing matrix.

$$\mathbf{y}_t = B\mathbf{x}_t \quad t = 1, 2, \ldots \tag{1}$$

By applying ICA to $\mathbf{x}_t$, we obtain the ICA components $\mathbf{y}_t$ of the image. Operating on these components leads to a codification algorithm. The codification approach is based on the idea that images with similar features may be restored from a common set of components. That is, it is possible to use ICA to define a set of basis functions to build a group of images. This set of basis functions is called the coding dictionary. It can be estimated using the FastICA algorithm, applied on image patches of various size [8].
We also have the demixing relation :

$$\mathbf{x}_t = W\mathbf{y}_t \quad t = 1, 2, \ldots \tag{2}$$

W is the demixing matrix, which can be deduced from the mixing matrix B.

3. WATERMARKING SCHEME

We have adopted the watermarking scheme proposed by Bounkong et al. [6]. We have adapted this method to be adequate for a video watermarking application, using two watermarking algorithms based respectively on quantization and trellis coding. In this section, we describe the elements of our approach.
In order to employ an ICA coding scheme, we have to build a coding dictionary adapted to the host data. This step is carried out using the popular FastICA algorithm, applied to image patches of size 16×16. The images, used for training, are selected from test video sequences, employed in our experiments (c.f. Section 4). Obtained data are centered and their dimensionality is reduced using principal component analysis. The preprocessed dataset are used as input for the FastICA algorithm.

3.1. Coefficients extraction

In order to preserve the quality of the watermarked video, we have chosen not to mark all frames, but to separate two marked frames by one distance. We have also chosen to insert the watermark in prediction error frames resulting from a motion compensation scheme. Therefore, a frame, selected to be marked, is firstly predicted from his previous neighbor. To carry out the prediction process, we use the motion compensation approach defined in MPEG-2 compression algorithm. The prediction error frame, P, is encoded using the ICA coding technique.
P is divided into patches of size 16×16, which represent a set of mixed signals. Each patch p_i is then demixed, using the demixing matrix W, resulting in a vector $\mathbf{x}_{p_i}$. A set of coefficients s_i is then randomly selected from $\mathbf{x}_{p_i}$, and watermarked in order to hide the message.
A mixing process applied on the watermarked vector allows to construct the marked frame.
We use two methods to watermark the coefficients s_i. The first one relies on a quantization system. The second one uses an informed trellis. In the following, both methods are described.

3.2. Quantization-based algorithm

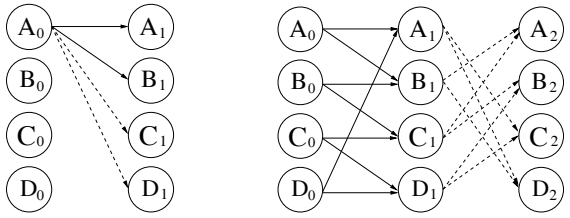
We extract m coefficients from each mixed vector $\mathbf{x}_{p_i}$, then we insert the same bit, in all of them, using a binary Quantization Index Modulation (QIM) scheme. In order to enhance the algorithm robustness, we embed the same bit in many mixed vectors. We will now give a description of the QIM scheme.
QIM is a blind watermarking technique where the host signal is quantized differently depending on the watermark information to be embedded [9]. A quantizer can be described by a set of reconstruction points Q in an L- dimensional space and a rule for assigning a length-L input signal to one of the points defined in Q. The basic principle of QIM can be described as follows:

- A set of different quantizers $\{Q_0, Q_1, Q_2, \ldots, Q_B\}$ is defined. The index set $B_s = \{0, 1, 2, \ldots, B\}$ denotes the B considered watermark messages.
- For embedding the watermark information $b \in B$, the host signal is quantized using the quantizer Q_b to obtain the public signal s. Thus, the expected embedding distortion D_E is equal to the introduced quantization noise.
- The watermark detector quantizes the received signal r by the union of all quantizers. The detector determines the index of the quantizer containing the reconstruction point closest to the received signal. This index corresponds to the received watermark information.

We call binary dither modulation the QIM scheme using a uniform scalar quantizer. In this case, we define two scalar quantizers Q_0, with step size Δ, and Q_1 with an offset of $\frac{\Delta}{2}$.

3.3. Watermarking using informed trellis

In this section, we present the second algorithm we have used to mark the error prediction frames. It is an informed spread spectrum technique: one message is represented by a set of vectors. At the coding stage, the most correlated vector with the host signal is selected to carry the message. The union of all sets constitutes the codebook. Error correcting codes, and trellis of convolutional codes in particular, provide practical way to construct and organize such a codebook. In this case, trellis is termed as informed trellis. On the other hand, our host signal is obtained by extracting n coefficients from each mixed vector $\mathbf{x}_{p_i}$.



(a) (b)

Fig. 1. A 4-state full connected trellis : In figure (a), we represent only the arcs issued from state A_0. Each bit can be coded by 2 different ways. In figure (b), we show how the full connected 4-state trellis can code a message of 2 bits.

We will now focus on the way we construct the informed trellis. Starting from a simple modification of a trellis code, we can produce an efficient code used for watermarking [10]. In a traditional trellis code, each possible message corresponds to a path through the trellis from a state A at time 0 to one of the nodes at time L. We refer to the transition from one column of nodes to the next column of nodes as a step, and each such step corresponds to one bit in the coded message. Each arc in the trellis is labelled with a randomly-generated length N vector. Each path, and thus each message, is coded with a length $L \times N$ vector that is the concatenation of the labels for the arcs it contains. This vector can be used as a watermark pattern. The trellis is modified so that multiple alternative code vectors can be obtained for each message. The basic idea is to have more than two arcs enter and exit each state, but still use each step of the trellis to encode a single bit. Thus, a given message can be represented by a number of different paths, and hence a number of different length $L \times N$ code vectors. We can increase the number of possible vectors for each message by allowing paths to start at any state at time 0.

Our trellis has 64 states and 64 arcs per state. Each arc is labelled with a vector of length $N = 128$. The label for each 1-arc is drawn from an independent, identically distributed Gaussian distribution. The label for each 0-arc is the negation of one of the 1-arc labels from the same node. The labels are scaled so that the mean squared error (MSE) between marked and unmarked images would equal the embedding strength, α.
In Figure 1, we present an example of a fully connected 4-state trellis, as well as the case when the trellis corresponds to a message of 2 bits.

4. RESULTS

In this section, we present the results obtained for both methods. Their performance was tested under MPEG-2 compression attacks.
We tested the methods on five 4:2:0 video sequences (Suzie, Hall monitor, Container, News, Mother and Daughter) composed of 150 frames in a QCIF format (176×144). In this format, pictures are composed of three components : luminance Y and color differences C_b and C_r, where the chrominance components are downsampled by a factor of 2 in both vertical and horizontal directions.
The watermarking distortion, determined by the mean PSNR value, was maintained at a level of 42 dB. We state now

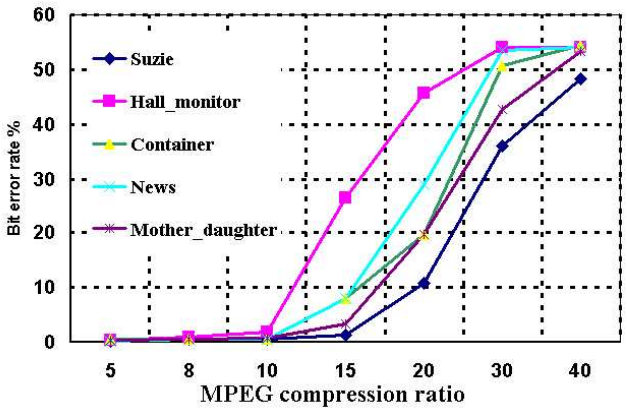


Fig. 2. Robustness of the method based on QIM scheme, against MPEG compression attack

parameters used in the quantization-based algorithm. We extracted 20 coefficients from a patch (c.f. Section 3.2, $m = 20$). In order to enhance robustness, the same bit was embedded in 9 patches chosen randomly. Thus, we were able to insert 11 bits in a frame of size 176×144. The quantizer step size, Δ, was set to 6. The MPEG compression attack results for this method are presented in figure 2

showing the percent of decoding bit errors as a function of the MPEG compression ratio. We can see that for a compression ration of 10, we were able to correctly decode most of bits. In the second approach, we extracted 41 coefficients

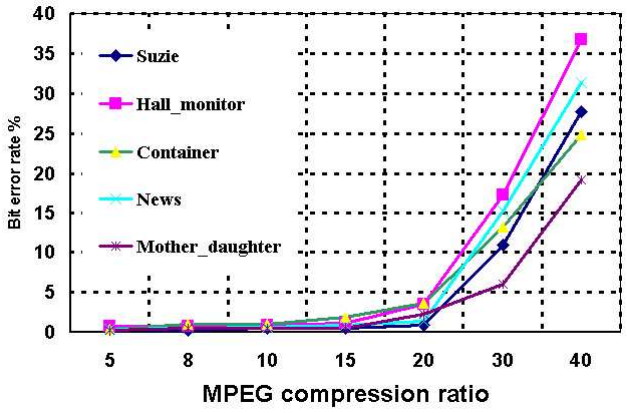


Fig. 3. Robustness of the method based on the informed trellis, against MPEG compression attack

from each patch (c.f. section 3.3, $n = 41$). This allowed us to embed 31 bits in a frame of size 176×144. The embedding strength α is equal to 1.5.
The robustness of this approach against MPEG compression attack is presented in figure 3. We retrieved most of embedded bits, after a compression ratio of 20.
In this paper, we do not adresse the temporal synchronisation problem of the watermarking scheme: if some attacks like frames elimination are applied to the watermarked sequence, it will result in a desynchronisation of the decoder. A synchronisation process could be established by considering a GOP-based decoder who would be synchronized in reference to the first frame in a GOP (Group Of Pictures).
Obviously, the second method is more robust against compression attack. Compared to quantization-based methods, spread spectrum watermarking methods are generally more robust against this kind of attack.
We performed a comparison test, in order to assess our choice to watermark prediction error frames. Therefore, we directly watermarked row images of the sequence. This led to a poor robustness level: obtained results point out that for an MPEG-2 compression ratio of order 8, the error bit rate exceeded 15%. This comparison test shows the improvement made by inserting watermark in error prediction frames: we were able to investigate the temporal information present in video sequences, in order to improve the scheme robustness.

5. CONCLUSION

In this paper, a video watermarking scheme is proposed. The insertion and extracting methods are performed in a set of statistically independent sources.
Two different ways, based respectively on quantization and trellis coding, are used to watermark the data.
Experimental results claim the invisibility and the robustness of the proposed approach: the scheme offers a good robustness against MPEG compression attack, as well as an important capacity level.

6. REFERENCES

[1] C. S. Lu, J. R. Chen, H. Y. M. Liao, and K. C. Fan, "Real-time MPEG-2 video watermarking in the VLC domain," in *International Conference on Pattern Recognition*, Quebec, Canada, August 2002.

[2] Y. Bodo, N. Laurent, and J. Dugelay, "Watermarking video, hierarchical embedding in motion vectors," in *IEEE International Conference on Image Processing*, Barcelona, Spain, September 2003.

[3] P. Bas, N. V. Boulgouris, F. D. Kovaros, J. M. Chassery, M. G. Strintzis, and B. Macq, "Robust watermarking of video object for MPEG-4 applications," in *SPIE Annual Meetting*, San-Diego, USA, 2001.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley-Interscience, 2001.

[5] F. J. González-Serrano, H. Y. Molina-Bulla, and J. J. Murillo-Fuentes, "Independent component analysis applied to digital watermarking," in *IEEE International Conference on Acoustic, Speech and Signal Processing*, Utah, USA, May 2001.

[6] S. Bounkong, B. Toch, D. Saad, and D. Lowe, "ICA for watermarking digital images," *Journal of Machine Learning Research*, vol. 4, pp. 1471–1498, 2003.

[7] A. S. Cohen and A. Lapidoth, "The gaussian watermarking game," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1639–1667, June 2002.

[8] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[9] B. Chen, *Design and Analysis of Digital Watermarking, Information Embedding and Data Hiding Systems*, Ph.D. thesis, Massachusetts Institute of Technology, 2000.

[10] M. L. Miller, G. J. Döer, and I. J. Cox, "Dirty-paper trellis codes for watermarking," in *IEEE International Conference on Image Processing*, New York, USA, September 2002.